Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

ON NEWTON'S METHOD OF APPROXIMATION

By Henry B. Fine

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY

Received by the Academy, August 8, 1916

In the first of the following theorems[1] a condition is given under which Newton's method of approximation for computing a real root of an equation $f(x)=0$ and the extension of this method used in computing a real solution of a system of equations $f_i(x_1, x_2, \ldots, x_n)=0$, $(i=1, 2, \ldots, n)$, will with certainty lead to such a root or solution. The condition relates to the absolute value of $f(x)$ or of the functions $f_i(x_1, x_2, \ldots, x_n)$ corresponding to the initial values of x and of the variables $x_1, x_2, \ldots, x_n$ respectively. No assumption is made as to the existence of a solution. On the contrary it is proved that under the condition to which reference has been made one and but one solution exists in a certain designated neighborhood of the initial x or $(x_1, x_2, \ldots, x_n)$. The second theorem is the extension of the first (for $n=1$) to the case of a complex root of an analytic equation $f(z) = 0$.

Theorem 1. *Let* $f_i(x_1, x_2, \ldots, x_n)$, $(i = 1, 2, \ldots, n)$, *be a system of real functions of the real variables* $x_1, x_2, \ldots, x_n$ *which have continuous first and second derivatives in the region* R, $(x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})$ *a set of values of* $x_1, x_2, \ldots, x_n$ *belonging to this region, and* $\xi_1, \xi_2, \ldots, \xi_n$ *the set of numbers determined by the equations.*

$$f_i\,(x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}) + \sum_{k=1}^{n} \frac{\partial f_i}{\partial x_k^{(0)}}\,\xi_k = 0 \qquad (i = 1, 2, \ldots, n)$$

and let S *denote the interval, circle, sphere, or hypersphere whose center is* $(x_1^{(0)} + \xi_1, \ldots, x_n^{(0)} + \xi_n)$ *and whose radius is* $\rho = \left(\sum_{k=1}^{n} \xi_k^2\right)^{\frac{1}{2}}$, S_ρ *being supposed to belong to* R.

Suppose also that in S *the functional determinant* F *of the functions* f_i *does not vanish,* $\mu(<\infty)$ *is the upper bound of the absolute values of the fractions whose denominators are* F *and whose numerators are the several first minors of* F, *and* $\nu(<\infty)$ *is the upper bound of the absolute values of the second derivatives of the functions* f_i.

Then, if

$$\left[\sum_{i=1}^{n} \{ f_i\,(x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}) \}^2\right]^{\frac{1}{2}} < \frac{1}{n^{7/2}\mu^2\nu} \qquad (a)$$

the equations $f_i = 0$ *have one and but one solution in* S, *and an approximate value of this solution, as close as may be desired, will be obtained by determinations first of* $\xi_1^{(j)}, \xi_2^{(j)}, \ldots, \xi_n^{(j)}$ *and then of* $x_1^{(j+1)}, x_2^{(j+1)}, \ldots, x_n^{(j+1)}$, *for* $j = 0, 1, 2 \ldots$ *successively, by the equations*

$$\begin{aligned} f_i\,(x_1^{(j)}, x_2^{(j)}, \ldots, x_n^{(j)}) + \sum_{k=1}^{n} \frac{\partial f_i}{\partial x_k^{(j)}}\,\xi_k^{(j)} &= 0, \qquad (i = 1, 2, \ldots, n) \\ x_k^{(j+1)} = x_k^{(j)} + \xi_k^{(j)}&, \qquad (k = 1, 2, \ldots, n), \end{aligned} \tag{b}$$

(*where* $\xi_k^{(0)} = \xi_k$), *the solution being* $\lim_{j=\infty} (x_1^{(j)}, x_2^{(j)}, \ldots, x_n^{(j)})$.

Moreover the equations (*b*) *will yield a similar sequence of approximations to this solution if instead of* $(x_1^{(0)}, x_2^{(0)}, \ldots x_n^{(0)})$ *any other point in* S *be chosen as the point of departure.*

For the case of a single equation $f(x) = 0$ *the condition* (*a*) *reduces to*

$$|\,f\,(x_0)\,| < \frac{\lambda^2}{\nu}, \tag{a'}$$

where λ (>0) *is the lower bound in the interval* S *of the values of* $|\,f'\,(x)\,|$.

For let

$$f_i^{(j)} = f_i\,(x_1^{(j)}, x_2^{(j)}, \ldots, x_n^{(j)}),\ \varphi_j = \left\{\sum_{i=1}^{n} [f_i^{(j)}]^2\right\}^{\frac{1}{2}},\ \rho_j = \left\{\sum_{k=1}^{n} [\xi_k^{(j)}]^2\right\}^{\frac{1}{2}}.$$

Then, if both $(x_1^{(j)}, x_2^{(j)}, \ldots, x_n^{(j)})$ and $(x_1^{(j+1)}, x_2^{(j+1)}, \ldots, x_n^{(j+1)})$ are in S,

$$f_i^{(j+1)} = f_i^{(j)} + \sum_{k=1}^{n} \frac{\partial f_i}{\partial x_k^{(j)}}\,\xi_k^{(j)} + \tfrac{1}{2} \sum_{k_1 l=1}^{n} \frac{\partial^2 f_i}{\partial x_k^{(j_i)}\,\partial x_l^{(j_i)}}\,\xi_k^{(j)}\,\xi_l^{(j)}, \quad (i = 1, 2, \ldots, n), \tag{1}$$

where each $x_k^{(j_i)}$ lies between $x_k^{(j)}$ and $x_k^{(j+1)}$.

Hence if the numbers $\xi_k^{(j)}$ be so taken as to satisfy the equations

$$f_i^{(j)} + \sum_{k=1}^{n} \frac{\partial f_i}{\partial x_k^{(j)}}\,\xi_k^{(j)} = 0, \quad (i = 1, 2, \ldots, n), \tag{2}$$

we shall have

$$f_i^{(j+1)} = \tfrac{1}{2} \sum_{k_1 l=1}^{n} \frac{\partial^2 f_i}{\partial x_k^{(j_i)}\,\partial x_l^{(j_i)}}\,\xi_k^{(j)}\,\xi_l^{(j)}, \quad (i = 1, 2, \ldots, n), \tag{3}$$

Solving the equations (2), we find

$$\xi_k^{(j)} = -\frac{\sum_{i=1}^{n} A_{ik}^{(j)} f_i^{(j)}}{F_j}, \quad (k = 1, 2, \ldots, n), \tag{4}$$

where F_j is the value of F for $(x_1, x_2, \ldots, x_n) = (x_1^{(j)}, x_2^{(j)}, \ldots x_n^{(j)})$, and $A_{ik}^{(j)}$ the cofactor of the element in the ith row and the kth column of F_j. But

$$\left| \sum_{i=1}^{n} A_{ik}^{(j)} f_i^{(j)} \right| \leqq \left\{ \sum_{i=1}^{n} [A_{ik}^{(j)}]^2 \right\}^{\frac{1}{2}} \left\{ \sum_{i=1}^{n} [f_i^{(j)}]^2 \right\}^{\frac{1}{2}}$$

Hence

$$\rho_i^2 \leqq \frac{\sum_{i,k=1}^{n} [A_{ik}^{(j)}]^2}{F_j^2} \sum_{i=1}^{n} [f_i^{(j)}]^2 \leqq n^2 \mu^2 \phi_j^2$$

and therefore

$$\rho_j \leqq n \, \mu \, \phi_j. \tag{5}$$

Again, the expressions under the summation sign in the equations (3) are of the form

$$\sum_{k,l=1}^{n} a_{kl} \, \xi_k \, \xi_l$$

and

$$\left| \sum_{k,l=1}^{n} a_{kl} \xi_k \xi_l \right| = \left| \sum_{k=1}^{n} \left(\sum_{l=1}^{n} a_{kl} \, \xi_l \right) \xi_k \right| \leqq \left\{ \sum_{k=1}^{n} \left(\sum_{l=1}^{n} a_{kl} \, \xi_l \right)^2 \right\}^{\frac{1}{2}} \left\{ \sum_{k=1}^{n} \xi_k^2 \right\}^{\frac{1}{2}}$$

$$\leqq \left(\sum_{k,l=1}^{n} a_{kl}^2 \right)^{\frac{1}{2}} \sum_{k=1}^{n} \xi_k^2 \leqq na \sum_{k=1}^{n} \xi_k^2, \tag{6}$$

where a denotes the greatest of the numbers $| a_{ik} |$.

Hence

$$\varphi_{j+1}^2 \leqq n \left(\frac{n\nu}{2} \right)^2 \left\{ \sum_{k=1}^{n} [\xi_k^{(j)}]^2 \right\}^{\frac{1}{2}} \leqq \frac{n^3 \, \nu^2}{2^2} \rho_i^4$$

and therefore

$$\varphi_{j+1} \leqq \frac{n^{3/2} \, \nu}{2} \rho_j^2. \tag{7}$$

From (5) and (7) it follows that

$$\varphi_{j+1} \leqq k_1 \, \phi_j^2, \text{ where } k_1 = \frac{n^{7/2} \mu^2 \nu}{2}, \tag{8}$$

$$\rho_{j+1} \leqq k_2 \, \rho_j^2, \text{ where } k_2 = \frac{k_1}{n\mu} = \frac{n^{5/2} \mu \nu}{2} \tag{9}$$

Hence $k_1 \varphi_{j+1} \leqq (k_1 \varphi_j)^2$ and therefore, by setting $j = 0, 1, 2, \ldots$ successively,

$$k_1 \varphi_j \leqq (k_1 \, \phi_0)^{2^j}. \tag{10}$$

Similarly

$$k_2 \rho_j \leqq (k_2 \rho)^{2^j}, \tag{11}$$

Let r denote any positive number less than 1, and suppose that $k_1 \varphi_0 < r$. Then since, by (5), $\rho \leqq n \mu \phi_0$, we also have $k_2 \rho < r$, and therefore, by (8) and (9),

$$k_2 \varphi_j < r^{2^j}, \qquad k_2 \rho_j r^{2^j}. \tag{12}$$

Again $\rho_{j+1} \leqq k_2 \rho_j \rho_j$ and therefore

$$\rho_{j+1} < r^{2^j} \rho_j. \tag{13}$$

Also, $k_2 \rho_j \leqq k_2 \rho \, (k_2 \rho)^{2^j-1}$ and therefore

$$\rho_j < r^{2^j-1} \rho. \tag{14}$$

Therefore, since $r < 1$, the sum $\rho_1 + \rho_2 \; . \; . \; . \; \rho_j$ increases with j to a limit which is less than $\rho r / (1 - r^2)$.

But $\rho r / (1 - r^2) < \rho$ when $r \leqq 1/2$.

Therefore when $r = 1/2$, that is when the condition $k_1 \varphi_0 < r$ becomes

$$\varphi_0 < \frac{1}{2k_1} = \frac{1}{n^{7/2} \mu^2 \nu}, \tag{15}$$

the point whose coordinates are $x_k^{(j)} = (x_k^{(0)} + \xi_k^{(0)}) + \xi_k^{(1)} + \; . \; . \; . + \xi_k^{(j-1)}$, $(k = 1, 2, \; . \; . \; . \; , n)$, will remain in S as j increases and will approach a definite limiting position $(c_1, c_2, \; . \; . \; . \; , c_n)$ in S as $j = \infty$. Moreover by (12), $\lim_{j=\infty} \varphi_j = 0$, and therefore, since the functions f_i are continuous, $f_i \; (c_1, c_2, \; . \; . \; . \; c_n) = 0$, $(i = 1, 2, \; . \; . \; . \; n)$. Therefore the equations $f_i = 0$ have the solution $(c_1, c_2, \; . \; . \; . \; c_n)$ in S.

Observe that it follows from (5) that when (15) is satisfied,

$$\rho < \frac{1}{2k_2} = \frac{1}{n^{5/2} \mu \nu} \tag{16}$$

The equations $f_i = 0$ have no other solution than $(c_1, c_2, \; . \; . \; . \; c_n)$ in S. For suppose that $(c_1 + h_1, \; . \; . \; . \; , c_n + h_n)$ represents another such solution. Then by developing as in (1) and applying (6)

$$0 = \sum_{k=1}^{n} \frac{\partial f_i}{\partial c_k} h_k + \frac{\theta_i}{2} n \nu \sum_{k=1}^{n} h_k^2, \quad (i = 1, 2, \; . \; . \; . \; , n), \tag{17}$$

where $| \theta_i | \leqq 1$.

Solving these equations for the numbers h_k in terms of $\sum_{k=1}^{n} h_k^2$, we obtain, by the method used in deriving (5) from (2),

$$\left[\sum_{k=1}^{n} h_k^2\right]^{\frac{1}{2}} \leqq \frac{n^{5/2}\mu\nu}{2}\sum_{k=1}^{n} h_k^2,$$

and therefore, if $\sum\limits_{k=1}^{n} h_k^2 \neq 0$,

$$\left[\sum_{k=1}^{n} h_k^2\right]^{\frac{1}{2}} \geqq \frac{2}{n^{5/2}\mu\nu} > 2\rho. \tag{18}$$

Hence $(c_1 + h_1, \ldots, c_n + h_n)$ lies outside of S. Therefore the existence of a second solution in S is impossible.

The equations (b) will also yield a sequence of approximations to $(c_1, c_2, \ldots, c_n)$ similar to the sequence $(x_1^{(j)}, x_2^{(j)}, \ldots x_n^{(j)})$ if instead of $(x_1^{(0)}, x_2^{(0)}, \ldots x_n^{(0)})$ any other point in S be taken as the point of departure. For let $(a_1^{(0)}, a_2^{(0)}, \ldots, a_n^{(0)})$ be any point in S, and $\xi_k^{(j)}$, $a_k^{(j+1)}$, $d_k^{(j)}$, $(k = 1, 2, \ldots, n)$, the numbers determined by the equations

$$f_i(a_1^{(j)}, a_2^{(j)}, \ldots, a_k^{(j)}) + \sum_{k=1}^{n} \frac{\partial f_i}{\partial a^{(j)}}\xi_k^{(j)} = 0, \quad (i = 1, 2, \ldots, n),$$

$$a_k^{(j+1)} = a_k^{(j)} + \xi_k^{(j)}, \quad d_k^{(j)} = c_k - a_k^{(j)}, \qquad (k = 1, 2, \ldots, n),$$

for $j = 0, 1, 2, \ldots$ successively.

Then since

$$c_k = a_k^{(j)} + d_k^{(j)} = a_k^{(j)} + (\xi_k^{(j)} + d_k^{(j+1)}), \quad (k = 1, 2, \ldots, n),$$

we shall have, on developing as in (1) and taking into account the equations which determine the numbers $\xi_k^{(j)}$,

$$0 = \sum_{k=1}^{n} \frac{\partial f_i}{\partial a_k^{(j)}} d_k^{(j+1)} + \frac{1}{2}\sum_{k,l=1}^{n} \frac{\partial^2 f_i}{\partial x_k^{(ji)}\partial x_l^{(ji)}} d_k^{(j)} d_l^{(j)}, \; (i = 1, 2, \ldots, n), \tag{19}$$

where each $x_k^{(ji)}$ lies between $a_k^{(j)}$ and c_k, $(k = 1, 2, \ldots n)$.

From (19) it follows, as in the proof of (5), that

$$\delta_{j+1} \leqq k_2\,\delta_j^2, \tag{20}$$

where $\delta_j = \left\{\sum\limits_{k=1}^{n} [d_k^{(j)}]^2\right\}^{\frac{1}{2}}$ and, as above, $k_2 = n^{5/2}\mu\nu/2$.

From (20) in turn it follows, as in (10), that

$$k_2\,\delta_j \leqq (k_2\,\delta_0)^{2^j}; \tag{21}$$

therefore, if $k_2\,\delta_0 < 1$, that $\lim_{j=\infty} \delta_j = 0$ and $\lim_{j=\infty} (a_1^{(j)}, a_2^{(j)}, \ldots, a_n^{(j)}) = (c_1, c_2, \ldots, c_n)$.

But δ_0 is the distance of the point $(a_1^{(0)}, a_2^{(0)}, \ldots, a_n^{(0)})$ from the point $(c_1, c_2, \ldots, c_n)$, and $k_2\delta_0 < 1$ if $\delta_0 < 1/k_2$; therefore, by (16), if

$\delta_0 < 2\rho$. And the condition $\delta_0 < 2\rho$ is satisfied if $(a_1^{(0)}, a_2^{(0)}, \ldots, a_n^{(0)})$ be taken anywhere in S.

Again since $k_2\,\delta_0 = r < 1$ we have, as in (12), $\delta_{j+1} \leqq (k_2\,\delta_j)\,\delta_j \leqq r^{2^j}\,\delta_j < \delta_j$. Hence each approximation in the sequence $(a_1^{(j)}, a_2^{(j)}, \ldots, a_n^{(j)})$, $(j = 0, 1, 2, \ldots)$, is closer than the one which precedes it.

Finally, by (21), $\delta_j \leqq r^{2^j}/k_2$. Let j' be any value of j such that $r^{2^{j'}}/k_2 < 0.1$ and $r^{2^{j'}} < 0.1$ and set $j = j' + l$ $(l = 0, 1, 2, \ldots)$. We then have

$$\delta_{j'+l} \leqq \frac{r^{2^{j'+l}}}{k_2} = \frac{r^{2^{j'}}}{k_2}\left(r^{2^{j'}}\right)^{2^l-1} < (0.1)^{2^l}.$$

Hence $\left\{\sum_{k=1}^{n} [a_k^{j'+l}]^2\right\}^{\frac{1}{2}}$ will coincide with $\left\{\sum_{k=1}^{n} c_k^2\right\}^{\frac{1}{2}}$ and therefore $a_k^{j'+l}$ with c_k $(k = 1, 2, \ldots, n)$, to at least the 2^lth decimal figure.

Theorem 2. *Let $f(z)$ be any function of the complex variable z which is analytic at $z = z_0$, and h the number determined by the equation*

$$f(z_0) + f'(z_0)\,h = 0.$$

Let S be the circular region whose center is $z_0 + h$ and whose radius is $\rho = |h|$, and suppose that, in S, $f(z)$ is everywhere analytic, $\lambda(>0)$ is the lower bound of the values of $|f'(z)|$ and $\nu(<\infty)$ is the upper bound of the values of $|f''(z)|$. Then, if

$$|f(z_0)| < \frac{\lambda^2}{\sqrt{2}\,\nu}, \qquad (a'')$$

the equation $f(z) = 0$ has one and but one root in S, and this root will be approximated to uninterruptedly by successive determinations of h_j and z_{j+1} by the formulas

$$f(z_j) + f'(z_j)\,h_j = 0, \quad z_{j+1} = z_j + h_j, \quad (j = 0, 1, 2, \ldots), \qquad (b')$$

(where $h_0 = h$), the root being $\lim_{j=\infty} z_j$.

For let

$$z_j = x + iy, \quad f(z_j) = \varphi(x, y) + i\,f(x, y), \quad h_j = \xi + i\eta.$$

Then if z_j and z_{j+1} are in S, we shall have

$$f(z_j + h_j) = f(z_j) + f'(z_j)\,h_j + \tfrac{1}{2}\left(\frac{\partial^2\varphi}{\partial x'^2}\,\xi^2 + 2\,\frac{\partial^2\varphi}{\partial x'\,\partial y'}\,\xi\eta + \frac{\partial^2\varphi}{\partial y'^2}\,\eta^2\right)$$
$$+ \frac{i}{2}\left(\frac{\partial^2\psi}{\partial x''^2}\,\xi^2 + 2\,\frac{\partial^2\psi}{\partial x''\,\partial y''}\,\xi\eta + \frac{\partial^2\psi}{\partial y''^2}\,\eta^2\right) \qquad (22)$$

where x', x'' lie between x and $x + \xi$; y', y'' between y and $y + \eta$. But

$$\left|\frac{\partial^2\varphi}{\partial x'^2}\xi^2+2\frac{\partial^2\varphi}{\partial x'\partial y'}\xi\eta+\frac{\partial^2\varphi}{\partial y'^2}\eta^2\right| = \left|\frac{\partial^2\varphi}{\partial x'^2}(\xi^2-\eta^2)-2\frac{\partial^2\psi}{\partial x'^2}\xi\eta\right|$$

$$\leqq\left\{\left[\frac{\partial^2\varphi}{\partial x'^2}\right]^2+\left[\frac{\partial^2\psi}{\partial x'^2}\right]^2\right\}^{\frac{1}{2}}\left\{(\xi^2-\eta^2)^2+4\xi^2\eta^2\right\}^{\frac{1}{2}}\leqq\nu\,|\,h_j\,|^2.$$

Similarly

$$\left|\frac{\partial^2\psi}{\partial x''^2}\xi^2+2\frac{\partial^2\psi}{\partial x''\partial y''}\xi\eta+\frac{\partial^2\psi}{\partial y''^2}\eta^2\right|\leqq\nu\,|\,h_j\,|^2.$$

Therefore

$$f(z_j+h_j)=f(z_j)+f'(z_j)\,h_j+\frac{\theta_1+i\theta_2}{2}\nu\,|\,h_j\,|^2,\quad -1<\theta_1,\theta_2<1,\quad (23)$$

Hence if h_j be so taken as to satisfy the equation $f(z_j)+f'(z_j)\,h_j=0$ and if, as in the proof of Theorem 1, we set $|\,h_j\,|=\rho_j$, $|\,f(z_j)\,|=\phi_j$, we shall have

$$\rho_j\leqq\frac{1}{\lambda}\varphi_j,\qquad \varphi_{j+1}\leqq\frac{\sqrt{2}\,\nu}{2}\rho_j^2.$$

But these are the same as the inequalities (5) and (7) in the proof of Theorem 1 for the case $n = 1$ except that ν is here replaced by $\sqrt{2}\nu$. It therefore at once follows from that proof, and in particular from the inequalities (15) and (18), that if $\varphi_0<\lambda^2/\sqrt{2}\nu$ and therefore $\rho=\rho_0<\lambda/\sqrt{2}\nu$, the equation $f(z)=0$ has one and but one root is S; also that the equations (b)' will yield a sequence of approximations to the root if instead of z_0 any other point in S be taken as the point of departure.[2]

---

[1] This paper was read before the American Mathematical Society, April 29, 1916.

[2] The distinctive feature of the method used in this paper is the determination of a number C such that if $\varphi_0<C$, then $\lim_{j=\infty}\varphi_j=0$. In this respect it differs from other discussions of Newton's method which I have been able to find. Of such discussions the following should be mentioned. Cauchy (Oeuvres [Ser. 2, Vol. 4], p. 573) obtained $\rho</\lambda 2\nu$ as the condition for the existence in S of a real root of a real equation obtainable by Newton's method, and $\rho<\lambda/4\nu$ as the corresponding condition for a complex root. Quite recently G. Faber (*J. Math., Berlin*, **138**) has proved, for an analytic equation $f(z)=0$, that if $|f(z)f''(z)/[f'(z)]^2|<\alpha<1$ in the circle whose center is z_0 and radius $\rho/(1-\alpha)$, a root of $f(z)=0$, obtainable by Newton's method, exists in this circle. I have been unable to discover any previous proofs of the existence of a solution for the case $n>1$. But Runge (*Encyc. Math. Wiss.*, **1**, p. 446) and E. Blutel (*Paris, Acad. Sci.*, C. *R.* **151**, 1109) have given proofs that if a solution exists at a certain point C, there must also exist a region R about C from any point of which a steady approach to C will be made by Newton's method Blutel's method of proof (an extension of one employed by Cauchy) being that which I have used in the proof of the corresponding part of Theorem 1.